Individual Differences in Aggression: Genetic Analyses by age, gender, and informant in 3-, 7-, and 10-year-old Dutch Twins

Hudziak, J.; van Beijsterveldt, C.E.M.; Bartels, M.; Rietveld, M.J.H.; Rettew, D.C.; Derks, E.M.; Boomsma, D.I.

published in
Behavior Genetics
2003

DOI (link to publisher)
10.1023/A:1025782918793

document version
Publisher's PDF, also known as Version of record

Behavior Genetics, Vol. 33, No. 5, September 2003 (© 2003)

Individual Differences in Aggression: Genetic Analyses by Age, Gender, and Informant in 3-, 7-, and 10-Year-Old Dutch Twins

J. J. Hudziak,[1,3] C. E. M. van Beijsterveldt,[2] M. Bartels,[2] M. J. H. Rietveld,[2] D. C. Rettew,[1] E. M. Derks,[2] and D. I. Boomsma[2]

Aggression in humans is associated with substantial morbidity and mortality. In this study we report on the aggressive behavior syndrome (AGG) in young children as defined by the Child Behavior Checklist (CBCL) and the Teacher Report Form (TRF). We assessed aggression in a large sample of Dutch twins at ages 3, 7, and 10 years. The purpose of this study was three-fold. First, we determined the number of children who are "clinically deviant" on the AGG scale. Second, we assessed the genetic and environmental contributions to AGG for the maternal, paternal, and teacher ratings at each age, for boys and girls. Third, we explored issues of rater bias by analyzing parental and teacher data simultaneously. CBCL data were available from mothers on 6436 three-year-old, 5451 seven-year-old, and 2972 ten-year-old twin pairs and CBCL data from fathers on 4207 three-year-old, 4269 seven-year-old, and 2295 ten-year-old twin pairs. Teacher report data from the TRF were collected for 1036 seven-year-old and 903 ten-year-old twin pairs from the Netherlands Twin Registry. Structural equation modeling was employed to obtain genetic and environmental estimates at each age. Analyses were conducted separately by age and informant, as well as simultaneously, for all informants. Differences in raw scores across gender were found, with boys being rated as more aggressive than girls by all informants. Mothers reported more symptoms than fathers, who reported more symptoms than teachers. Evidence for moderate to high genetic influence (51%–72%) was seen for AGG by all three informants at all ages with only small sex differences in heritability estimates. Best fitting models for AGG by parent reports also included a small contribution of common environment. The largest sex differences in heritabilities were seen at age 10. Contributions of common (13%–27%) and unique (16%–31%) environment were small to moderate. There was some evidence of genetic dominance by teacher report for 10-year-old girls.

KEY WORDS: Aggression; heritability; parents; teachers; twins; childhood.

INTRODUCTION

Human aggression can be described in many ways. The purpose of this paper is to report on aggressive behavior as defined by the aggressive Syndrome (AGG) of the Child Behavior Checklist (CBCL; Achenbach, 1991a) and Teacher Report Form (TRF; Achenbach, 1991c). The AGG syndrome (see Table I for a complete description of AGG items) is associated with negative behavioral outcomes and has been shown to be a good predictor of DSM oppositional defiant disorder (ODD) in young children and conduct disorder (CD) in older children. Loeber *et al.* (2000) describe the essential features of ODD as "a recurrent pattern of negativistic, defiant, disobedient, and hostile behavior toward authority figures, which leads to impairment, and the essential features of CD are a repetitive and persistent pattern of behavior in which the basic rights of others and major age-appropriate societal norms or

[1] Departments of Psychiatry and Medicine, (Division of Human Genetics), Center for Children Youth and Families, University of Vermont, College of Medicine, Burlington Vermont.

[2] Department of Biological Psychology, Vrije Universiteit, Van der Boechortstraat 1, 1081 BT Amsterdam, The Netherlands.

[3] To whom correspondence should be addressed at Division of Child Psychiatry and Behavioral Genetics, University of Vermont, College of Medicine, 1 South Prospect, Burlington Vermont 05405. Tel: 802-656-1084. e-mail: jhudziak@zoo.uvm.edu

0001-8244/03/0900-0575/0 © 2003 Plenum Publishing Corporation

Table I. CBCL-Items of the Syndrome Aggressive Behavior for the CBCL/2-3 and CBCL/4-18

CBCL/2-3 (9 items)	CBCL/4-18 (20 items)	TRF items (25)
Cruel to animals	Argues a lot	Argues a lot
Destroy own things	Bragging, boasting	Bragging, boasting
Destroy others' things	Cruelty, bullying, or meanness to others	Cruelty, bullying, or meanness
Disobedient	Demands a lot of attention	Defiant, talks back to staff
Gets in many fights	Destroys his/her own things	Demands a lot of attention
Hits other	Destroys things belonging to his/her family or others	Destroys others' things
Hurts animals or people	Disobedient at home	Destroys own things
Physically attacks people	Disobedient at school	Disturbs other people
Unusually loud	Easily jealous	Disobedient at school
	Gets teased a lot	Easily jealous
	Physically attacks people	Gets in many fights
	Screams a lot	Talks out of turn
	Showing off or clowning	Physically attacks people
	Stubborn, sullen or irritable	Disrupts classroom discipline
	Sudden changes in mood or feelings	Screams a lot
	Talks to much	Showing off or clowning
	Teases a lot	Explosive and unpredictable behavior
	Temper tantrums or hot temper	Easily frustrated, must have demands met
	Threatens people	Stubborn, sullen, or irritable
	Unusually loud	Sudden changes in mood or feelings
		Talks too much
		Teases a lot
		Temper tantrums or hot behavior
		Threatens people
		Unusually loud

Note: DSM-IV ODD ITEMS:
1. Often loses temper
2. Often argues with adults
3. Often actively defies or refuses to comply with adults' requests or rules
4. Often deliberately annoys people
5. Often blames others for his or her mistakes or misbehavior
6. Is touchy or easily annoyed by others
7. Is often angry and resentful
8. Is often spiteful and vindictive

rules are violated." It is estimated that approximately 50% of youth who meet criteria for CD will later meet criteria for antisocial personality disorder (ASPD; Burke *et al.,* 2000). Several studies (Biederman *et al.,* 1993; Chen *et al.,* 1994; Steingard *et al.,* 1992) have reported that the AGG syndrome is not only a strong marker of DSM-ODD, it is also useful as an indicator of CD. In a U.S. sample, 100% of children with a *T* score >67 (which corresponds to a raw score on AGG in the top 5% of the distribution) on AGG fulfilled DSM-III-R ODD criteria (Chen *et al.,* 1994). In a family study, we tested AGG as a predictor of DSM-IV ODD and found AGG to be a highly sensitive and specific predictor in both probands and their siblings with ODD (Hudziak *et al.,* 2003).

AGG in childhood appears to be a stable characteristic that persists into adulthood. For example, Verhulst and Koot (1991) studied the stability of parent-reported behaviors in a Dutch longitudinal sample of 4- to 16-year-old children. Correlations for aggressive behavior ranged from .55 to .67 for 2-, 4-, and 6-year age intervals. AGG raw scores are higher in boys than girls and decrease with age in both genders (e.g., see Stanger *et al.,* l997, and Figure 1). AGG raw scores also differ by informant, with mothers rating more AGG than fathers who rate more symptoms than teachers (Achenbach, 1991a, 1991b).

Many, if not most, studies to date on the prevalence and heritability of AGG have been carried out in samples of children of different ages, raising the possibility that age differences may affect the results. One of the aims of our work is to determine the raw scores of AGG by age, gender, and informant to determine if any of these factors affect the genetic and environmental influences on AGG. Each is discussed in some detail.

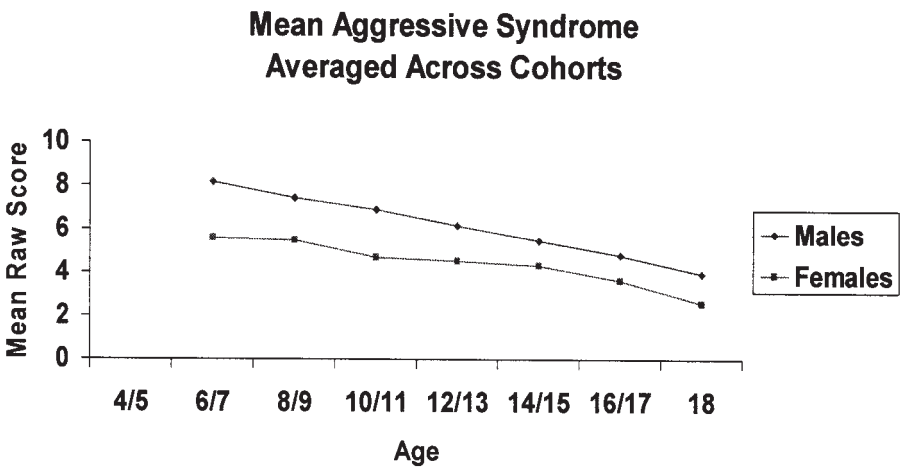

Fig. 1. Means aggressive syndrome scores averaged across cohorts, separately for males and females. Figure adapted from Stanger *et al.* (1997). Figure appears in Stanger, C., Achenbach, T. M., & Verhulst, F. C. (1997). Accelerated longitudinal comparisons of aggressive versus delinquent behavior syndromes. *Development and Psychopathology, 9,* 43–58.

Genetic Influences on Human Aggressive Behavior

There is evidence that AGG, as measured from parental report data by the CBCL, is a heritable trait. Edelbrock *et al.* (1995) reported genetic variance of 50% for AGG and shared environmental variance of 15%. Ghodsian-Carpey and Baker (1987) reported a heritability coefficient (h^2) of 94% for AGG. van den Oord *et al.* (1996) found both genetic (69%) and shared environmental influences (12%) on AGG in one of the first studies from the Netherlands Twin Registry. The analyses were based on data from over 1300 Dutch 3-year-old Dutch twin pairs. For the broad-band Externalizing (EXT) Scale of the CBCL, which includes the AGG syndrome, van der Valk *et al.* (1998, 2001, 2003) reported also genetic and shared environmental influences for EXT in Dutch twin pairs at ages 3 and 7 years. At age 3, a significant positive sibling interaction for externalizing behavior was found, indicating cooperation between two children of a twin pair. Further, analyses with partly overlapping samples at ages 10 and 12 years indicated significant influences of genetic and shared environmental factors (Bartels *et al.,* 2003a, b). Hudziak *et al.* (2000) in a study of twins from age 8–12 years in a U.S. twin population reported a heritability of 61% for males and 69% for females for AGG. Others, using different measures of aggression have also argued for genetic influences of aggressive behavior (Bergeman and Seroczynski, 1998). In a recent article by Vierikko *et al.* (2003), aggressive behavior was measured in a large sample of 11- to 12-year-old twin pairs. The aggression measure was a six-item scale derived from the Multidimensional Peer Nomination Inventory as reported by parents and teachers. Vierikko *et al.* (2003) reported significant genetic and common environmental influences on aggression. In addition, sex-specific findings were reported that vary by informant. For both parental and teacher ratings, boys showed lower levels of heritability and higher levels of shared environment than girls. In addition, the teacher rating data also suggested the presence of either sex-specific genetic effects (heterogeneity) or sex-specific shared environmental effects. In a meta-analysis, Miles and Carey (1997) reported an overall genetic effect that accounted for half of the variance in aggression. In addition, they reported heritability estimates that varied by age. In the young (l8 years or younger), both genes and common environment were important, but for adults genetic influences were identified as the primary influence on AGG behavior. Rhee and Waldman (2002) conducted a meta-analysis to estimate the heritability on antisocial behavior and reported age differences. For children (below age 13), adolescents (age 13–18), and adults (above age 18), both genetic and common environmental factors were important, but the magnitude of the environmental influences decreased as age increased. Finally, there are reports of relations of candidate genes to aggressive behavior. Schmidt *et al.* (2002), reported that children with the long repeat allele of the DRD4 gene had significantly higher scores on CBCL AGG at age 4 than children with the short alleles. Caspi *et al.* (2002) reported on a functional polymorphism of the MAOA gene in interaction with childhood maltreatment as it relates to the development of violent aggressive behavior at adult age. In this publication, maltreatment was reported to confer greater risk for antisocial behavior, providing evidence for gene-environment

interaction. Brunner *et al.* (1993) reported an association between a null allele at the MAOA locus that was related to antisocial behavior in a large Dutch pedigree.

Development and Aggression

Aggressive behavior is common in humans and typically diminishes with age (Figure 1). Therefore when attempting to study AGG, it is important to take into account developmental findings. For example, a maternal report of a raw score of 8 for a boy at age 6 would be considered as falling in the statistically "normal" range, where the same score at age 12 would yield a score that falls in the "clinical range." AGG is not the only behavior that changes over time. Delinquent behavior scale scores tend to first decrease then increase over the course of normal development (Stanger *et al.,* 1997). In this paper we report on analyses of data that control for age in twins studied at successive developmental periods. The onset of antisocial behaviors such as AGG has been the focus of a considerable amount of research (e.g., Moffitt, 1993). The analyses reported in this paper are cross-sectional rather than longitudinal, but the different age-groups allow the study of genetic and environmental contributions to AGG from the very young (age 3) into childhood (age 10).

Gender and Aggression

Figure 1 is also instructive in that it demonstrates differences in the raw scores of AGG between boys and girls at all ages. Gender plays a key role when deviance on behavioral rating scales is estimated. For instance, a raw score of 7 for a boy at age 10 is in the normal range, but the same score would be in the deviant range for a girl of the same age. Such differences in raw scores between genders must be taken into account when analyzing behavioral data. Genetic studies of childhood psychopathology should allow for potential gender differences in the manifestation of genotypes. According to Hartung and Widiger (1998), of the 21 disorders usually first diagnosed before adulthood, 17 have higher prevalence in boys than girls. They enumerated several sources of error that could generate or exaggerate gender differences in rates of psychopathology, most notably sampling biases and biases within the diagnostic criteria, concluding that "there may not be a mental disorder for which there are not important gender differences in the manner in which the disorder is expressed." (p. 274). Nowhere is this truer than in the study of childhood aggression and oppositional conduct disorder/conduct disorder (ODD/CD). Boys are more often diagnosed with ODD/CD (Last, 1989; Lewis and Miller, 1990). In this study we will estimate the prevalence of AGG by analyzing *T* scores that have been standardized by gender, by three different informants (mothers, fathers, and teachers) to determine if clinically deviant AGG is more common in males than females. Once accomplished, we will then analyze these data using gender-genetic models to determine if there are gender differences in the genetic architecture of AGG.

Informants and Aggressive Behavior

When studying emotional/behavioral disorders in children, researchers often depend on parents as informants. Parents usually only moderately agree on how to best describe their child's behavior. On average the correlations between mother and father reports on CBCL syndromes is .59, but the interparent agreement on reports of AGG is often higher. In fact, the correlations of AGG in the CBCL/2-3 were .71 (Achenbach, 1992). Mothers report higher raw scores than fathers. Boys typically have higher scores than girls (Achenbach and Rescorla, 200l) by all informants.

Although agreement is higher with measures of AGG than other syndromes, it remains true that differences exist in the way in which mothers and fathers report on their children. Informant differences are even more robust when comparing parent reports to teacher reports. The cross-informant correlations between teachers and parents for reports of AGG are consistently in the 0.38 or lower range (Achenbach and Rescorla, 2001). One reasonable way to consider the lack of agreement is that each informant is commenting on different behaviors based on their relationships with the child. For instance, mothers report what mothers see, which may be different from a father's or teacher's perspective. Another source of rater disagreement can be rater bias. This bias (Neale and Stevenson, 1989; Hewitt *et al.,* 1992) could arise if the parents' own traits influenced ratings (a projection bias), or if parents exhibited response biases (e.g., stereotyping, employing different normative standards, or having certain response styles, i.e., judging problem behaviors more or less severely). Further, in studying the etiology of childhood psychopathology using twin pairs and parental/teacher ratings, contrast effects could be of significance (Carey, 1986; Eaves, 1976). The importance of contrast effects has been studied mainly for ADHD and attention problems (e.g., Simonoff *et al.,* 1998; Rietveld *et al.,* 2003a) and is suggested for

traits that show low to very low DZ correlations in addition to variance differences between MZ and DZ twins. Such trends in the data are typically not seen for AGG, and contrast effects thus do not seem to play an important role.

The results for aggression suggest an influence of common family environment. We investigate if the magnitude of these common environmental influences on AGG changes by informant. Specifically, we test models that estimate rater bias, because the phenotype that different raters agree on usually tends to be less influenced by common environmental contributions. Our datasets with AGG rated by mothers, fathers, and teacher offer the opportunity to disentangle sources of rater agreement and disagreement. These analyses are important to perform, because if rater bias exists, and is not tested for, the resulting findings may include an overestimation of shared environmental influences on AGG. Therefore we fit a psychometric rater model on the mother, father, and teacher reports at age 3, 7, and 10 years. Previous studies of externalizing problems on the significance and magnitude of rater bias and rater-specific views showed that rater bias accounted for at most 15% of the total variance. Four percent to 10% of the total variance was accounted for by the rater-specific views. After correcting for an overestimation of shared environment by rater bias, significant but small influences of shared environmental factors on externalizing behavior at age 3,7, 10, and 12 were still found (Bartels *et al.,* 2003a, b; van der Valk, 2001, 2003).

In summary, studies on AGG in general population samples reveal it to be a common and stable condition. In clinical studies, AGG has been shown to be predictive of DSM ODD and CD, which in turn are highly predictive of adult antisocial behavior (ASPD; Robins and Rutter, 1990). In twin studies, AGG has been shown to be heritable and also influenced by common environmental factors (Rhee and Waldman, 2002). ODD/CD and ASPD in turn have been shown to be influenced by both genetic and environmental factors and associated with profoundly negative adult behaviors and outcomes. Twin studies of AGG can be interpreted as shedding light on the genetic contributions to a developmental syndrome that often precedes more severe forms of adult psychopathology, including, but not limited to, antisocial personality disorder.

This article comments on the genetic and environmental influences on AGG as measured across a severity continuum. The genetic and environmental estimates presented in this report comment on nonpathological as well as pathological levels of AGG. We provide distributions of raw scores of the sample, by age, gender, and informant that indicate what percentage of the sample falls in the statistically deviant range, and therefore, are at risk for ODD, CD, and ASPD.

This is the first report that we know of that presents cross-sectional analyses of AGG on twins when they were 3, 7, and 10 years old. These analyses are unique in that they provide cross-sectional and rater-bias models of different informants (mothers, fathers, and teachers) on the same children.

METHOD

Subjects and Procedure

The data of the present study are derived from a large ongoing longitudinal study, which examines the genetic and environmental influences on the development of problem behavior in families with 3- to l2-year-old twins. The families are volunteer members of the Netherlands Twin Register, kept by the Department of Biological Psychology at the Vrije Universiteit in Amsterdam (Boomsma, 1998; Boomsma *et al.,* 2002). From 1986 families with twins are recruited a few months after birth. Currently, 40%–50% of all multiple births are registered by the Netherlands Twin Registry. For the present study, we included data of 3- and 7-years-old twin pairs from cohorts 1986–1994, and of 10-year-old twin pairs from cohorts 1986–1991. Parents of twins were asked to fill in questionnaires about problem behavior for the eldest and youngest twin at ages 3, 7, and 10 years. After 2 months a reminder was sent to the nonresponders, and after 4 months those who still did not respond were telephoned. This procedure resulted in a response rate (at least one questionnaire was returned) of 77% for age 3. From ages 3 to 7, and ages 7 to 10 the continued participation was 80%. At ages 7 (available for cohort 1992–1994) and 10 (available for cohort 1989–1991), teachers were asked to fill in the Teacher Report Form. Permission to approach teachers was obtained from the parents. After 2 months a reminder was sent to the nonresponders.

For 822 same sex twin pairs zygosity was based on blood group polymorphisms ($n = 424$) or DNA ($n = 398$). For the remaining twins, zygosity was determined by questionnaire items, filled by the mother, about physical similarity and frequency of confusion of the twins by family and strangers (Goldsmith, 1991). The classification of zygosity was based on a discriminant analysis, relating the questionnaire items to zygosity based on blood/DNA typing in a group of same-sex twin pairs. The zygosity was correctly

Table II. Number of Twin Pairs for Parental and Teacher Reports

	Age 3		Age 7			Age 10		
	Mother	Father	Mother	Father	Teacher	Mother	Father	Teacher
MZM	1045	685	922	733	181	525	418	153
DZM	1057	710	894	704	160	470	346	140
MZF	1213	788	1064	828	214	617	479	202
DZF	990	649	852	655	151	455	359	125
DOSm_f	2131	1375	1719	1349	330	905	693	283
Total	6436	4207	5451	4269	1036	2972	2295	903

classified by questionnaire in nearly 95% of the cases (Rietveld *et al.,* 2000).

A family was excluded when one of the twin pair had a disease or handicap that interfered severely with normal daily functioning (about 2%). Table II gives an overview of the number of families with complete twin pairs. Socioeconomic status (SES) was obtained from a full description of the occupation of the parents when the children were 3 years of age. The level of occupation was coded according to the system used by Statistics Netherlands (CBS, 1993). The code was based on the mental complexity of the work and ranged from low skilled to scientific work. An earlier comparison of the parental SES distribution with those obtained for the general Dutch population showed a slightly higher frequency of the middle and higher SES groups (for details see Rietveld *et al.,* 2003b).

Attrition rates as well as a detailed discussion on the representativeness of the sample at each age are discussed in detail elsewhere in this special issue (van Beijsterveldt *et al.,* 2003).

Measures

At age 3, behavior was measured with the CBCL/2-3, a questionnaire that included 100 items that describe specific behavioral, emotional, and social problems (Achenbach and Rescorla, 2000). Parents were asked to rate the behavior that the child displayed now or within the past 2 months on a 3-point scale: 0 if the problem item was not true, 1 if the items were somewhat or sometimes true, and 2 if it was very true or often true. The syndrome aggressive behavior is constructed for the Dutch population (see (Koot *et al.,* l997)) and is comparable with the syndrome scale as developed by Achenbach (1992).

At ages 7 and 10 years problem behavior was measured with the CBCL/4-18 (Achenbach, 1991a), a questionnaire of 118 items developed to measure problem behavior in children 4–18 years old. Again parents were asked to rate the behavior of the child of the preceding 6 months on a 3-point scale. The aggressive behavior syndrome contains items that partially overlap with the CBCL/2-3. For the CBCL/4-18 eight syndrome scales were composed according to the 1991 profile (Achenbach, 1991a, b). In the present study, subjects with more than three missing items were not included in the analyses. This occurred in less than 2.5% of the received questionnaires filled out by the mother when the twins were aged 7 and 10 years.

At ages 7 and 10 teachers, using the Teacher's Report Form (TRF), also rated problem behavior (Achenbach, 1991c). The psychometric properties of the TRF, like those of the CBCL are described in detail elsewhere (Achenbach and Rescorla, 2001). Teachers were requested to rate behavioral problems as 0, 1, or 2, like the rating of the CBCL, based on a 2-month period. The items for both the CBCL and the TRF are presented in Table I.

Statistical Analyses

The generalized Linear modeling procedure (GLM) (SPSS version 10.0 for Windows) was used to test for mean differences between raters (father and mother), gender, and ages. For each age, the AGG scores of father and mother reports were analyzed as the dependent variables, with sex as the between-subjects factor and rater as the within-subjects factor. To test for age differences (only between ages 7 and 10), the within-subjects factor age was added into the model. For the teacher report, in the GLM the AGG scores at age 7 and 10 were used as dependent variables, and sex and age (age 7 and 10 were independent groups) as between-subjects factors. Analyses were conducted separately for the oldest and youngest of the twin pair, because due to their genetic relatedness their data are not independent. Means and standard deviations were estimated using Mx (Neale *et al.,* 1999). PRELIS 2 was used to obtain the variance-covariance

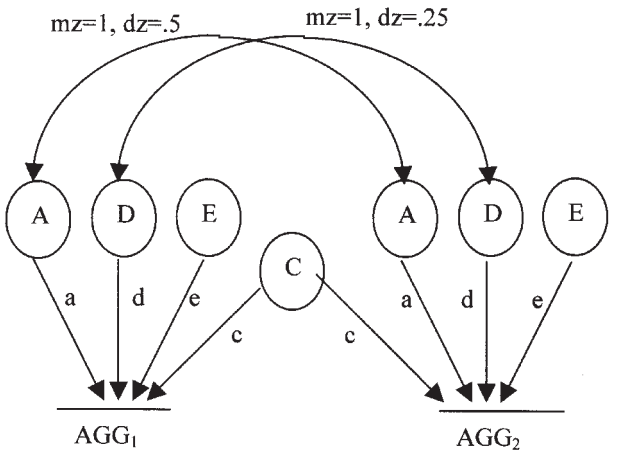

Fig. 2. Univariate path model. *Note:* A = Additive genetic; D = dominance genetic; E = unique environment; C = shared environment; AGG_1 = aggression problem score of twin 1; AGG_2 = aggression problem score of twin 2.

matrices of the observations, separately for each age-sex-by-zygosity group. These matrices were used as input for the single-informant genetic analyses. Structural equation modeling was employed to obtain estimates of the genetic and environmental contributions to the observed variances and covariances between measures. These analyses were conducted in Mx, using maximum likelihood estimation of parameters. The confidence intervals for the estimated genetic and environmental parameters were also obtained from Mx. Technical details of genetic model-fitting analyses are reviewed elsewhere (Neale and Cardon, 1992). Figure 2 summarizes the fundamental univariate genetic model that underlies these analyses. This model was used to estimate the additive genetic (A, additive effects of genes at multiple loci), dominance genetic (D, interaction of genetic effects at the same loci), common or shared environment (C, shared among members of the same household), or nonshared environment (E, unique to the individual) effects. The circles represent the latent, unmeasured factors. Correlations between latent factors were 1.0 for MZ and 0.5 (for A) and .25 (D) for DZ pairs. The environment shared by two members of a pair (C) is assumed not to depend on the zygosity of the twins. The unique or nonshared environment is by definition, uncorrelated between two members of a pair. Estimates of the unique environmental effects also include measurement error (Eaves, 1982; Plomin *et al.,* 1997).

First, an ADE or ACE model with parameters allowed to differ between genders was fitted to the data. Different models were evaluated on the basis of their goodness-of-fit (chi-squared statistic). The associated degrees of freedom equal the number of statistics minus the number of estimated parameters. The significance of the C or D was tested by dropping this parameter. Sex differences were tested by constraining the parameters to be equal across gender. When sex differences appeared to be significant, a scalar sex-limitation model was tested. In this model a difference in total variance between boys and girls is allowed, but the relative contributions of genetic and environmental influences are the same for boys and girls. Comparison of models was also done by the use of Akaike's Information Criterion (AIC; Akaike, l987), which is a goodness-of-fit index that takes parsimony (i.e., number of parameters) into account. The AIC is the χ^2 minus two times the degrees of freedom.

Psychometric Rater Models

To gain insight into the magnitude of rater-specific views and rater bias and to get an impression of how rater bias might affect the results, a second series of analyses were carried out. Based on prior studies (Bartels *et al.,* 2003a, b; van der Valk *et al.,* 2001, 2003), we used the psychometric model to disentangle sources of rater agreement and disagreement. For a full description of the psychometric model see Bartels *et al.* (2003c, this issue). In short, the variance-covariance matrix of AGG rated by mothers, fathers, and teachers for twin 1 and twin 2 can be decomposed into a matrix of genetic, shared environmental, and nonshared environmental variances for behavior similarly assessed by all raters and for parts of the behavior on which raters disagree. Disagreement between raters in this model can be caused by rater-specific behavioral views, leading to different but valid information of each rater. The psychometric model tests this possibility by examining whether there are significant genetic effects on the rater-specific part of each rater's rating. If the behaviors uniquely rated by the parents or teacher are shown to be influenced by the genotype of the child, the rater must have been assessing a "real" but unique aspect of the child's behavior. Error and unreliability cannot cause the systematic effects necessary for the model to estimate these genetic influences. Disagreement can also be caused by rater bias, which will confound the rater-specific shared environmental effects. Rater bias that represents for instance certain response styles, stereotyping, or the use of specific normative standards will be estimated in the rater-specific shared environmental factors, because the rater will bias the ratings of both children of a twin pair in the same way. To

make optimal use of all available data, these analyses were carried out on the raw data.

RESULTS

Table II describes the number of twin pairs, by gender, zygosity, and age. The sample size decreases by age by virtue of the fact that this is an ongoing longitudinal study in which we add newborn twins annually. The sample size decreases by informant, with mothers providing more responses than fathers and fathers more than teachers. Questionnaires from fathers were not collected for a period of around 3 years, and we only recently began collecting teacher data. At all ages and by all informants, boys were reported to have higher raw scores on AGG than girls. Table III presents the calculated prevalence for deviance on AGG, by informant and gender, at each age in two ways. First we applied the cut-points as Koot *et al.* (1997) did, to determine what percentage of our twin sample would be considered deviant by the standard set forth from analyses of a sample of nonreferred Dutch children. The Koot *et al.* cut-points were developed by determining a raw score cut-off that identifies the top 5% of scores on AGG as deviant (e.g., fell in 95th percentile or greater this corresponds to a *T* score of 67). In addition to simply applying the cut-points from the Dutch normative sample, we also computed *T* scores on our twin population. We did this for two reasons. One to provide a basis for comparison, and second, because each of our samples are larger than the original Dutch normative sample, we wanted to compute *T* scores for each informant, for each age, for each gender, rather than compute *T* scores averaged across ages 4–11 and 12–18, as were reported by Koot *et al.* (1997). We would expect that our samples would have somewhat higher rates of clinical deviance than a nonreferred population because in our twin sample we did not exclude "cases."

In the case of the 3-year-old boys and girls, roughly 8% of the boys and girls exceed both the Koot/Verhulst and the Twin Sample community borderline cut-off of a *T* score of greater than 67. In this instance, it did not matter whether we used the nonreferred community standard or the twin standard cut-points. Although it may be difficult to distinguish true problematic aggressive behavior of 3-year-olds from age-related manifestations of normal development, these distributions yield evidence that approximately 8% of the children are rated as deviant on the AGG scale. In comparison with other reports on pre-school rates of psychopathology and related behavior problems, these percentages appear to be consistent to quite modest (Campbell and Cueva, 1995; Koot *et al.,* 1997). The data for the 7- and 10-year-old age-groups indicate that between 4%–9% of boys and of girls exceeded a *T* score of 67 on reports of AGG, and thus may be at risk for DSM-ODD/CD. Because of our much larger

Table III. Prevalence of Aggressive Behavior ($T \geq 67$) for Mother, Father, and Teacher Report at Ages 3, 7, and 10 Years

				Prevalence aggression (Koot *et al.*)		Prevalence aggression (Twin sample)	
	N	♂	♀	♂	♀	♂	♀
Aggm3	13131	6453	6678	523 (8.1%)	516 (7.7%)	523 (8.1%)	516 (7.7%)
Aggf3	8623	4275	4348	246 (5.8%)	260 (6.0%)	246 (5.8%)	260 (6.0%)
Aggm7	11062	5428	5634	403 (7.4%)	488 (8.7%)	337 (6.2%)	397 (7.0%)
Aggf7	8651	4278	4373	209 (4.9%)	271 (6.2%)	266 (6.2%)	344 (7.9%)
Aggt7	2215	1098	1117	67 (6.1%)	70 (6.3%)	87 (7.9%)	77 (6.9%)
Aggm10	6037	2934	3103	191 (6.5%)	196 (6.3%)	191 (6.5%)	196 (6.3%)
Aggf10	4647	2247	2400	83 (3.7%)	98 (4.1%)	185 (8.2%)	177 (7.4%)
Aggt10	1946	942	1004	71 (7.5%)	61 (6.1%)	71 (7.5%)	61 (6.1%)

sample, we are able to compute age- and gender-specific cut-points, which provide a more consistent finding, that equal numbers of girls as boys meet criteria for clinical deviance on AGG. To put this in a clinical context, Biederman *et al.* (1993), Steingard *et al.* (1992), and Chen *et al.* (1994), have all reported that children who exceed a *T* score of 67 on AGG are likely to meet criteria for DSM-III-R ODD. Given these data, equal numbers of girls and boys would be at risk for DSM-ODD based on their AGG scores, yet six times more boys than girls are diagnosed with DSM-ODD (Loeber *et al.,* 2000). These data raise the possibility that the lack of gender normative data for DSM ODD symptoms lead to the under diagnosis, and thus treatment, of ODD in girls.

Table IV describes the means and variances on the measures of AGG by age and gender for Maternal (M), Paternal (P), and Teacher (T) reports. Lower total mean scores on the 3 year-old twin sets are affected by the differences in the total number of items on the 2/3 CBCL AGG syndrome (9) versus the total number of items on the 4–18 CBCL (20). As also found by Stanger *et al.* (1997) as children age from 7 to 10 there is a reduction in the mean scores on AGG behavior [$F(1,1864) = 96.623$, $p < 0.001$]. At all ages, mothers report more AGG symptoms than fathers for both boys and girls. The differences in mean scores between mother and father reports across all ages were significant [age 3 years: $F(1,4068) = 104{,}440$, $p < 0.001$; age 7: $F(1,4237) = 148{,}632.613$, $p < 0.001$; age 10: $F(1,2257) = 98.613$, $p < 0.001$]. In addition, all informants reported more AGG in boys than in girls. The difference in means was significant across all ages [age 3: $F(1,4068) = 171.288$, $p < 0.001$; age 7: $F(1,4237) = 114.468$, $p < 0.001$; age 10: $F(1,2257) = 43.746$, $p < 0.001$]. For teacher reports there were no significant age effects [$F(1,2076) = .579$, $p = .447$], but there were significant gender effects [$F(1,2073) = 135.719$, $p < 0.001$], with boys scoring higher than girls. Note that the results of the GLMs are given only for the oldest twin, but the same results were obtained for the youngest twin.

It should be reiterated that the 3-year-old CBCL AGG scale is different than the scale for 7- and 10-year-olds. Unfortunately, direct comparisons via standardization of scores based simply on the number of items cannot be done because the items themselves emerged from factor analyses that indicated that many of the AGG items in the CBCL/4-18 simply are not reported often enough for 2- to 3-year-old children to be included in factor analytic results. Therefore, although we could standardize based simply on the number of items, it would not do justice to the fact that not only the number of items change across development so does the magnitude of the factor loadings. For boys and girls, we noted an increase in AGG from age 3 to age 7, but given the differences in the number of items on the CBCL/2-3 (10 items) and the CBCL/4-18 (20 items) we did not compute for statistical significance here. It is unclear what differences account for how teachers vs. mothers and fathers report on their children's AGG behavior; however, a number of explanations have been suggested. These include the different setting the child is in during the rating period, differences in interaction between rater and subject, and differences in the TRF AGG scale construction.

Twin correlations are provided in Table V. MZ twins show high correlations by all informants. These are approximately .70 by TRF for all age-groups and approximately .80 by M(other) and F(ather) by all age-groups. The size of these correlations were stable for both genders by M and F reports at ages 3, 7, and 10 and for T reports at ages 7 and 10. This suggests that parents and teachers continue to report AGG behavior in a fairly reliable way, even in the face of diminished raw scores across development, and across measures (CBCL/TRF). An estimation of the influence of the environment unique to the individual is obtained after subtracting the monozygotic correlation from unity. This calculation suggests that around one fifth of the

Table IV. Means and Variances (between brackets) of the Transformed AGG Score

	MZM			DZM			MZF			DZF			DOS_{males}			$DOS_{females}$		
	Mother	Father	Teacher	Mother	Father	Teacher	Mother	Father	Teacher	Mother	Father	Teacher	Mother	Father	Teacher	Mother	Father	Teacher
Age 3	1.89	1.75	–	1.79	1.71	–	1.47	1.34	–	1.42	1.28	–	1.72	1.59	–	1.31	1.26	–
	(0.74)	(0.75)	–	(0.72)	(0.69)	–	(0.72)	(0.69)	–	(0.72)	(0.72)	–	(0.71)	(0.73)	–	(0.67)	(0.66)	–
Age 7	2.62	2.45	1.85	2.49	2.31	1.86	2.13	1.90	1.00	2.10	1.91	1.12	2.42	2.21	1.80	1.98	1.80	1.00
	(1.39)	(1.38)	(2.65)	(1.41)	(1.36)	(2.67)	(1.47)	(1.45)	(1.62)	(1.47)	(1.38)	(1.52)	(1.46)	(1.34)	(2.36)	(1.42)	(1.30)	(1.50)
Age 10	2.45	2.19	1.93	2.33	2.10	1.79	1.94	1.67	1.02	1.97	1.77	1.16	2.27	2.04	1.94	1.85	1.63	.98
	(1.63)	(1.53)	(2.97)	(1.57)	(1.39)	(2.87)	(1.31)	(1.34)	(1.71)	(1.37)	(1.25)	(1.65)	(1.56)	(1.54)	(2.60)	(1.32)	(1.26)	(1.48)

Table V. Twin Correlations for the CBCL AGG Score as Reported by Parents and Teacher

	MZM			DZM			MZF			DZF			DOSm_f			r(M,F)		r(T,M)		r(T,F)	
	M	F	T	M	F	T	M	F	T	M	F	T	M	F	T	boys	girls	boys	girls	boys	girls
Age 3	.81	.78	–	.55	.51	–	.83	.83	–	.53	.50	–	.47	.49	–	.68	.65	–	–	–	–
Age 7	.83	.84	.72	.49	.51	.33	.84	.84	.71	.53	.58	.33	.51	.50	.26	.73	.73	.33	.26	.26	.27
Age 10	.83	.84	.73	.50	.53	.41	.79	.81	.73	.54	.50	.25	.47	.53	.17	.74	.71	.38	.30	.39	.27

Note: r(M,F) = Interparent correlation; r(T,M) = correlation between teacher and mother rating; r(T,F) = correlation between teacher and mother rating. In the last four columns the interparent correlations and correlation between teacher and mother ratings is given.

total variance for AGG by M and F reports and one quarter by T reports of the total variance AGG is explained by the unique environment. DZ correlations are moderate and stable for all groups by M and F report. By TRF the DZ correlations are smaller. DZ correlations are greater than half of MZ correlations at all agegroups by all informants, except TRFs on females at ages 7 and 10 years. Thus, if at all, one would expect genetic dominance to be important only for TRFs reports on females. The final columns of Table V give the across M and T correlations on AGG at ages 7 and 10. Correlations on M and T report range from .33 to .38 (for boys) and from .26 to .30 (for girls) across ages 7 and 10. The correlations between F and T are comparable with the M and T correlations. These acrossinformant correlations are consistent with what Achenbach (1991c) reported for the CBCL-TRF correlations in a U.S. normative sample and what Verhulst and Koot (1991) reported for a Dutch normative sample. The data on the lower correlations between mothers and teachers on measures of AGG in girls is particularly interesting. We have found a similar result in a study of mother and teacher reports on DSM-IV ODD symptoms. Teachers report fewer symptoms than parents and teacher's reports on boys more closely resemble parental reports than they do for girls. The data may provide a window on why more boys than girls are referred for clinical care for the AGG/ODD behavioral problems than girls. Simply put, it appears that girls are able to modify their AGG/ODD behavior in school settings better than boys.

Genetic Modeling

We fitted a series of models that tested for the relative contribution of additive genetic (A), dominant

Table VI. Model Fitting Results for CBCL Aggression as Rated by Mother and Father

		Mother			Father		
	df	χ^2	AIC	p	χ^2	AIC	p
Age 3							
ACE sex differences	9	15.45	−2.55	0.08	**7.91**	**−10.09**	**0.54**
AE sex differences	11	91.79	69.79	0	53.55	31.55	0
ACE no sex differences	12	**21.79**	**−2.21**	**0.04**	30.13	6.13	0
AE no sex differences	13	94.813	68.81	0.0	73.27	47.27	0
ACE scalar	11	–	–	–	21.11	−.088	0.03
Age 7							
ACE sex differences	9	6.31	−11.69	0.71	6.87	−11.13	0.65
AE sex differences	11	53.83	31.83	0	70.04	48.04	0
ACE no sex differences	12	**7.04**	**−16.97**	**0.86**	**12.92**	**−11.08**	**0.38**
AE no sex differences	13	53.87	27.87	0	70.11	44.11	0
Age 10							
ACE sex differences	9	**4.99**	**−13.00**	**0.83**	8.90	−9.10	0.45
AE sex differences	11	31.76	9.76	0	41.37	19.37	0.06
ACE no sex differences	12	35.42	11.42	0	21.54	−2.46	0.04
AE no sex differences	13	58.29	32.29	0	52.64	26.64	0
ACE scalar	11	12.208	−9.79	0.35	**13.97**	**−8.03**	**0.24**

Note: **Bold** models were the best fitting models.

Table VII. Model Fitting Results for the CBCL Aggression as Reported by the Teacher

	df	χ^2	AIC	p
Age 7				
ADE sex differences	9	6.900	−11.10	0.648
ACE sex differences	9	9.677	−8.323	0.377
AE sex differences	11	9.739	−12.261	0.554
AE no sex differences	13	83.778	57.778	0.0
AE scalar	**12**	**9.781**	**−14.22**	**0.635**
Age 10				
ADE sex	9	11.168	−6.832	0.264
ACE sex	9	16.935	−1.065	0.05
ACE boys, ADE girls	9	10.883	−7.117	0.284
AE boys; ADE girls	**10**	**11.168**	**−8.832**	**0.345**
AE sex	11	20.724	−1.276	0.036
AE scalar	12	21.091	−2.909	0.049

Note: **Bold** models were the best fitting models.

genetic (D) or common environment (C), and unique environmental (E) influences. Note that estimating D and C at the same time is not possible in a design using only twin MZ and DZ twins reared together. In each model we tested for differences across gender. When sex differences appeared, a scalar sex-limitation model was tested. The data are presented by informant, age, and model. For each analysis, only competing or best fitting models are presented. Hence, for analyses of M and F data, we present ACE and AE models with and without gender differences. The best fitting models for the parent ratings were ACE models with and without gender differences (see Table VI). This was true for all ages and both raters. A scalar sex-limitation model, in which the total variance was allowed to differ between boys and girls, gave a better fit to the data for F at age 10. A slightly different set of candidate models emerged for the T analyses. The best fitting models for T reports varied by age (see Table VII). For 7-year-old pairs an AE-scalar model, while for 10-year-olds an AE model for boys and an ADE model for girls were found.

The estimates of the additive genetic, dominant genetic, shared, and unique environmental contributions to the total variance are summarized in Table VIII. Additive genetic influences accounted for 51%–71% of the variance, shared environmental influences for 12%–27%, and unique environmental influences for 16%–23% of the variance for reports on males at ages 3, 7, and 10. Genetic influences for boys were lowest (51% by father) at age 3. Genetic and environmental influences were relatively stable across development for boys by both parent and teacher reports. For girls, the additive genetic influences accounted for between 52% and 69% of the variance, shared environment for between 15% and 27% of the variance, and unique environment for between 16% and 31% of the influence. Over ages, the magnitude of genetic and environmental influences on AGG behavior are relatively stable by M and F report.

Our final model aimed to determine the percentage of AGG that is similarly assessed by all raters versus the aspects that are unique to each rater. The results are summarized in Table IX. When AGG similarly assessed by all raters is taken into account, the heritability is in the range of 60%–79%. This heritability estimate is quite stable over gender and age.

Table VIII. For the Parental and Teacher Data, Estimates of Relative Contributions of Genetic and Environmental Factors

	Mother			Father			Teacher		
	a^2	c^2	e^2	a^2	c^2	e^2	a^2	d^2	e^2
Age 3									
Boys	61% (56%–65%)	21% (16%–24%)	18% (17%–20%)	51% (41%–62%)	26% (16%–36%)	23% (20%–25%)			
Girls				67% (59%–74%)	15% (09%–23%)	17% (15%–19%)			
Age 7									
Boys	66% (61%–71%)	18% (13%–22%)	16% (15%–17%)	61% (55%–66%)	22% (17%–27%)	17% (16%–18%)	69% (64%–74%)	–	31% (26%–36%)
Girls									
Age 10									
Boys	71% (60%–78%)	12% (5%–23%)	17% (15%–19%)	58% (51%–66%)	23% (16%–30%)	19% (17%–20%)	72% (65%–78%)	–	27% (22%–34%)
Girls	52% (40%–67%)	27% (13%–39%)	21% (19%–24%)				21% (3%–51%)	49% (19%–69%)	30% (24%–36%)

Note: The 95% confidence interval is given between the brackets.

Table IX. Percentage of Total Variance of AGG on Which All Raters Agree Accounted for by Additive Genetic, Shared Environmental, and Nonshared Environmental Influences (left part of the table) and the Percentage of the Total Variance of AGG (rater agreement and disagreement) Accounted for by Rater-Specific A, C, and E

	Rater agreement			Rater disagreement								
				Rater specific A			Rater specific C			Rater specific E		
	Additive genetic	Shared environment	Unique environment	M	F	T	M	F	T	M	F	T
Age 3												
Boys	.60	.24	.16	.14	.12	–	.10	.08	–	.10	.08	–
Girls	.78	.11	.10	.16	.16	–	.08	.09	–	.11	.10	–
Age 7												
Boys	.79	.07	.14	.14	.09	.51	.09	.11	.00	.07	.08	.19
Girls	.72	.15	.12	.13	.04	.29	.11	.15	.13	.07	.08	.20
Age 10												
Boys	.77	.09	.14	.19	.01	.32	.03	.17	.16	.07	.06	.20
Girls	.74	.10	.16	.07	.09	.47	.16	.13	.00	.09	.08	.18

Idiosyncratic experiences (without measurement error) explain 10%–15% of the behavior similarly assessed by all raters. If the results of the rater analyses are compared to the analyses per rater (Table VIII), the expected change in shared environmental influences, and accordingly change in heritability, is observed. However, as can be seen in Table IX, rater bias, defined as rater-specific C accounts for 16% of the total variance at most. As expected, teachers provide additional information on the child's behavior, represented by relatively large estimates for rater-specific A for teachers.

DISCUSSION

In this cross-sectional study of AGG behavior in twins at ages 3, 7, and 10 we have demonstrated a number of important findings. First we have replicated the findings of Stanger *et al.* (1997) that ratings of aggression decrease over the 7-year developmental period encompassed by our data base. Second, we have confirmed that boys are consistently rated as more aggressive than girls at all ages by all informants. Further we demonstrated that mothers report more symptoms of AGG than fathers, who report more symptoms than teachers. This conforms to the reports of Verhulst and Koot (1991) in their general epidemiological study.

We demonstrate that AGG is as common in girls as it in boys for children ages 3 to 10 years. Our data indicate that clinically deviant scores on AGG are more common than on other behavioral problems such as attention problems (Rietveld *et al.,* 2003b). Between 4% and 9% of boys and girls in this study were reported to exceed a *T* score of 67 by mother, father, or teacher report. These data allow the reader to get a sense of the prevalence and the severity associated with AGG behavior. As stated earlier, others have repeatedly demonstrated that a *T* score of greater than 67 on AGG predicts DSM ODD and CD. Given the findings that AGG predicts ODD/CD, which, in turn, predicts risk for substance use disorder and adult antisocial personality disorder, these findings represent significant evidence for addressing AGG as a health care problem.

A second remarkable finding is the consistent magnitude of the gender/genetic effects on AGG. Despite the variance seen in raw scores across gender, age, and informants, the magnitude of the genetic and environmental contribution to AGG is relatively stable. Even though raw scores decrease as a child ages, the magnitude of the genetic influences on these behaviors remains about the same. In a related paper in this issue (van Beijsterveldt *et al.,* 2003, this issue) longitudinal genetic models are fit to test whether or not the genetic and environmental influences identified here at each successive age are due to the same or new genetic or environmental factors. These data, taken together, indicate that AGG is a highly heritable behavior, with genetic effects explaining the majority of the influence at each age.

Despite the differences in informants, the lack of correlation often seen among maternal, paternal, and teacher reports, genetic modeling results in general agreement about the magnitude and nature of the genetic and environmental contributions to AGG behavior. The two sources of information, that is, parental and teacher ratings, also led to some differences. Analyses of maternal and paternal CBCL reports provide consistent evidence of additive genetic and

shared and unique environmental influences across development. Teacher reports provide no evidence of shared environment. Teacher reports on males indicate high additive genetic (70%) and modest unique environmental contributions to AGG behavior (30%). The results agree with an adoptive sibling study that examined CBCL AGG in middle childhood, which reported genetic influences on AGG for parents and teacher ratings. Shared environmental influences were only found for parent's ratings of AGG (Deater-Deckard and Plomin, 1999). Reasons for these rater differences may be that the TRF and CBCL AGG measure not the same behavior or that both type of raters observe the child in different situations, possibly with their own biases. This is in agreement with the results of the psychometric rater model that showed that the teacher ratings contained a substantial individual view component.

In addition to providing data based on multiple raters separately, insights into the etiology of rater disagreement is presented as well. The majority of the influence on AGG, either common to all raters or rater specific is additive genetic. The influence of C was either absent or, when present, quite modest. Rater bias accounts for 10%–15% of the total variance at most. The use of multiple raters is recommended, especially when considering the additional information a teacher provides is taken into account. So although father and mother provide significant (but small) rater-specific data, teacher data provide much more specific information than is provided by mother or father alone or together. Therefore in studies on the genetics of AGG it will be important to include teacher data to obtain a full picture of AGG.

These studies support further molecular genetic investigations into AGG and antisocial behavior by providing evidence of genetic influences on AGG. Given the robust genetic influences reported here at each age by all informants, it will be important to pursue approaches that aim to determine which genes contribute to these very serious behavioral problems. Given the Caspi *et al.* (2002) findings that environmental factors interact with genotypes in the development of antisocial behavior, studies that aim to determine the molecular genetic contributions to AGG should consider directly measuring environmental contributions to phenotypic identification. It will be important to consider environmental factors by gender, given our findings on gender differences. Such work could lead to a change in how AGG and antisocial behaviors are diagnosed, perceived, and treated. The notion that ODD/CD is intractable must be questioned to make advances in therapeutics (Burke *et al.,* 2000). Research that considers discrete genetic and environmental contributors to these behaviors may identify targets for treatment.

Taken together, our data indicate AGG to be highly heritable across ages, gender, and informant. Further, estimates of prevalence of the extreme phenotype indicate that AGG and its corollary DSM diagnoses (ODD/CD) are quite common. Although the direct relations between high AGG scores and DSM-ODD were not directly tested in this study, multiple prior studies have provided evidence that children who fall in the top 5% of the AGG scoring distribution are at risk for ODD. Thus these data indicate that AGG and possibly ODD represent a highly prevalent and therefore important health care problem. These relations are the focus of an ongoing study in which we combine CBCL and DSM assessments in large numbers of twin pairs.

When interpreting these results it will be important for the reader to remember that this longitudinal study was carried out in a normal population, rather than a clinical population. First, the analyses of AGG on these twins at different ages are not independent of each other, as these are the same twins, just studied at different ages. This, of course, is one of the main points of our design and this report. We wanted to study the same children in order to be assured if there were major changes in the magnitudes of genetic or environmental influences on AGG so we could be assured that it was not accounted for by having different children in the samples at different ages. Second, AGG was measured across the behavioral spectrum, with some of the twins having very low scores on AGG and others high scores, with the majority in between. The stability and change data presented here are group data and do not reflect individual cases. Thus children who are low on AGG at one time point are likely to be low at a second, similarly a child who is high at one time point is expected to be high at a second time point, given the stability data presented. Whether or not this is true for a specific child is not addressed in this paper. Finally, it is clear that high scores on the AGG scale are associated with significant psychopathology (Hudziak *et al.,* in press) such as ODD and CD and later antisocial personality disorder (ASPD). It is less clear, however, what moderate or low scores on the AGG scale might mean. Moderate scores on AGG may actually be beneficial to the child, whereas extremely low scores may be associated with unfavorable behavioral traits of anxiety and avoidance. Thus one should be careful when interpreting this data for individual cases. Finally, in all cases we find moderate environmental influences on

AGG, and thus the argument that all AGG behavior is genetic is not made or supported by these reports.

ACKNOWLEDGMENTS

This research is supported by NIH grant RO1 MH58799-03, to J. Hudziak and Spinoza and NWO 575-25-006 grants to D. Boomsma.

REFERENCES

Achenbach, T. M. (1991a). *Integrative guide for the 1991 CBCL/4-18, YSR, and TRF profiles*. Department of Psychiatry, Burlington, VT: University of Vermont.

Achenbach, T. M. (1991b). *Manual for the Child Behavior Checklist/4-18*. University of Vermont, Burlington, VT: Department of Psychiatry.

Achenbach, T. M. (1991c). *Manual for the Teacher's Report Form and 1991 Profile*. University of Vermont, Burlington, VT: Department of Psychiatry.

Achenbach, T. M. (1992). *Manual for the Child Behavior Checklist/2-3 and 1992 Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M., and Rescorla, L. A. (2000). *Manual for the ASEBA Preschool Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children Youth, & Families.

Achenbach, T. M., and Rescorla, L. A. (2001). *Manual for the ASEBA School-Age Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children Youth, & Families.

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika* **52**:317–322.

Bartels, M., Boomsma, D. I., Rietveld, M. J. H., van Beijsterveldt, C. E. M., Hudziak, J. J., and van den Oord, E. J. C. G. (2003a). Disentangling genetic, environmental, and rater effects on internalizing and externalizing problem behavior in 10-year-old twins. *J. Child Psychology and Psychiatry* (under revision).

Bartels, M., Hudziak, J. J., Boomsma, D. I., Rietveld, M. J. H., van Beijsterveldt, C. E. M., and van den Oord, E. J. C. G. (2003b). A study of parent ratings of internalizing and externalizing problem behavior in 12-year-old twins. *J. Am. Acad. Child. Adolesc. Psychiatry* (in press).

Bartels, M., Hudziak, J. J., van den Oord, E. J. C. G., van Beijsterveldt, C. E. M., Rietveld, M. J. H., and Boomsma, D. I. (2003). Co-occurrence of aggressive behavior and rule-breaking behavior at age 12: Multirater analyses. *Behavior Genetics* **33**:607–622.

Bergeman, C. S., and Seroczynski, A. D. (1998). Genetic and environmental influences on aggression and impulsivity. In M. Maes, and E. F. Coccaro (Eds.), *Neurobiology and clinical views on aggression and impulsivity* (Vol. 5); (pp. 63–80). Chichester: John Wiley and Sons.

Biederman, J., Faraone, S. V., Doyle, A., Lehman-Krifcher, B., Kraus, I., Perrin, J., and Tsuang, M. T. (1993). Convergence of the Child Behavior Checklist with structured interview-based psychiatric diagnoses of ADHD children with and without comorbidity. *J. Child Psychol. Psychiatry* **34**:1241–1251.

Boomsma, D. I. (1998). Twin registers in Europe: An overview. *Twin Res.* **1**:34–51.

Boomsma, D. I., Vink, J. M., van Beijsterveldt, C. E. M., de Geus, E. J. C., Beem, A. L., Mulder, E. J. C. M., Riese, H., Willemsen, A. H. M., Bartels, M., van den Berg, M., Derks, E. M., de Graaff, S. C., Kupper, H. M., Polderman, J. C., Posthuma, D. I., Rietveld, M. J. H., Stubbe, J. H., Knol, L. I., Stroet, T., and van Baal, G. C. M. (2002). Netherlands Twin Register. *Twin Res.* **5**:401–406.

Brunner, H. G., Nelen, M., Breakfield, X. O., Ropers, H. H., and Oost, B. A. (1993). Abnormal behavior associated with a point mutation in the structural gene for monoamine oxidase A. *Science* **262**:578–580.

Burke, J. D., Loeber, R., and Birmaher, B. (2000). Oppositional defiant disorder and conduct disorder: A review of the past 10 years: II. *J. Am. Acad. Child Adolesc. Psychiatry* **39**:1275–1293.

Campbell, M., and Cueva, J. E. (1995). Psychopharmacology in child and adolescent psychiatry: A review of the past seven years: II. *J. Am. Acad. Child Adolesc. Psychiatry* **34**:1262–1272.

Carey, G. (1986). Sibling imitation and contrast effects. *Behav. Genet.* **16**:319–341.

Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., Taylor, A., and Poulton, R. (2002). Role of genotype in the cycle of violence in maltreated children. *Science* **297**:851–854.

CBS. (1993). *Standaard Beroepenclassificatie 1992* [standardized classifications of occupations 1992]. Voorburg/Heerlen, The Netherlands: Central Bureau of Statistics.

Chen, W. J., Faraone, S. V., Biederman, J., and Tsuang, M. T. (1994). Diagnostic accuracy of the Child Behavior Checklist scales for attention deficit-hyperactivity disorder: A receiver-operating characteristic analysis. *J. Consult. Clin. Psychol.* **62**:1017–1025.

Deater-Deckard, K., and Plomin, R. (1999). An adoption study of the etiology of teacher and parent reports of externalizing behavior problems in middle childhood. *Child Dev.* **70**:144–154.

Eaves, L. J. (1976). A model for sibling effects in man. *Heredity* **36**:205–214.

Eaves, L. J. (1982). The utility of twins. In V. E. Anderson, W. A. Hauser, J. K. Penry, and C. F. Sing (Eds.), *Genetic basis of the epilepsies* (pp. 249–276). New York: Raven Press.

Edelbrock, C., Rende, R., Plomin, R., and Thomson, L. A. (1995). A twin study of competence and problem behavior in childhood and early adolescence. *J. Child Psychol. Psychiatry* **36**:775–785.

Ghodsian-Carpey, M., and Baker, L. A. (1987). Genetic and environmental influences on aggression in 4- to 7-year old twins. *Aggress. Behav.* **13**:173–186.

Goldsmith, H. H. (1991). A zygosity questionnaire for young twins: A research note. *Behav. Genet.* **21**:257–270.

Hartung, C. M., and Widiger, T. A. (1998). Gender differences in the diagnosis of mental disorders: Conclusions and controversies of the DSM-IV. *Psychol. Bull.* **123**:260–278.

Hewitt, J. K., Silberg, M. C., Neale, M. C., Eaves, L. J., and Erickson, M. (1992). The analysis of parental ratings of children's behavior using LISREL. *Behav. Genet.* **22**:293–317.

Hudziak, J. J., Copeland, W., Stanger, C., and Wadsworth, M. (2003). Diagnosing externalizing behavior problems with the child behavior checklist: A quality receiver-operating curve analysis. *J. Child Psychology and Psychiatry and Allied Disciplines* (in press).

Hudziak, J. J., Rudiger, L., Neale, M., Heath, A., and Todd, R. A. (2000). A twin study of inattentive, aggressive, and anxious/depressed behavior. *J. Am. Acad. Child Adolesc. Psychiatry* **39**:469–476.

Koot, H. M., van den Oord, E. J. C. G., Verhulst, F. C., and Boomsma, D. I. (1997). Behavioral and emotional problems in young preschoolers: Cross-cultural testing of the validity of the Child Behavior Checklist/2-3. *J. Abnorm. Child Psychol.* **25**:183–196.

Last, C. G. (1989). Anxiety disorders. In T. H. Ollendick and M. Hersen (Eds.), *Handbook of child psychopathology* (pp. 219–228). New York: Plenum Press.

Lewis, M., and Miller, S. M. (1990). *Handbook of developmental psychopathology*. New York: Plenum Press.

Loeber, R., Burke, J., Lahey, B., Winters, A., and Zera, M. (2000). Oppositional defiant and conduct disorder: A review of the past 10 years: I. *J. Am. Acad. Child Adolesc. Psychiatry* **39**:1468–1484.

Miles, D. R., and Carey, G. (1997). Genetic and environmental architecture of human aggression. *J. Pers. Social Psychol.* **72**:207–217.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychol. Rev.* **100**:674–701.

Neale, M. C., and Stevenson, J. (1989). Rater bias in the EASI Temperament scales: A twin study. *J. Pers. Soc. Psychol.* **56**:446–455.

Neale, M. C., and Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Dordrecht: Kluwer Academic.

Neale, M. C., Boker, S. M., Xie, G., and Maes, H. H. (1999). *Mx: Statistical Modeling* (5th ed.). VCU Box 900126, Richmond, VA 23298: Dept. of Psychiatry (5th ed.).

Plomin, R., DeFries, J. C., McClearn, G. E., and Rutter, M. (1997). *Behavioral genetics*. New York: W. H. Freeman.

Rhee, S. H., and Waldman, I. D. (2002). Genetic and environmental influences on antisocial behavior: A meta-analysis of twin and adoption studies. *Psychol. Bull.* **128**:490–529.

Rietveld, M. J., Hudziak, J. J., Bartels, M., van Beijsterveldt, C. E., Boomsma, D. I. (2003a). Heritability of attention problems in children. I: Cross sectional results from a study of twins, age 3–12 years. *Am. J. Med. Genet.* **117B**:102–112.

Rietveld, M. J. H., Posthuma, D., Dolan, C. V., and Boomsma, D. I. (2003b). ADHD: Sibling interaction or dominance—An evaluation of statistical power. *Behav. Genet.* **33**:247–255.

Rietveld, M. J. H., VanderValk, J. C., Bongers, I. L., Stroet, T. M., Slagboom, P. E., and Boomsma, D. I. (2000). Zygosity diagnosis in young twins by parental report. *Twin Res.* **3**:134–141.

Robins, L., and Rutter, M. (1990). *Straight and devious pathways from childhood to adulthood.* Cambridge: Cambridge University Press.

Schmidt, L., Fox, N. A., Rubin, K. H., Hu, S., and Hamer, D. H. (2002). Molecular genetics of shyness and aggression in preschoolers. *Pers. Indiv. Diff.* **33**:227–238.

Simonoff, E., Pickles, A., Hervas, A., Silberg, J. L., Rutter, M., and Eaves, L. (1998). Genetic influences on childhood hyperactivity: Contrast effects imply parental rating bias, not sibling interaction. *Psychol. Med.* **28**:825–837.

Stanger, C., Achenbach, T. M., and Verhulst, F. C. (1997). Accelerated longitudinal comparisons of aggressive versus delinquent syndromes. *Dev. Psychopath.* **9**:43–58.

Steingard, R., Biederman, J., Doyle, A., and Sprich-Buckminster, S. (1992). Psychiatric comorbidity in attention deficit disorder: Impact on the interpretation of the Child Behavior Checklist results. *J. Am. Acad. Child Adolesc. Psychiatry* **31**:449–454.

van Beijsterveldt, C. E. M., Bartels, M., Hudziak, J. J., and Boomsma, D. I. (2003). Causes of stability of aggression from early childhood to adolescence: A longitudinal genetic analysis in dutch twins. *Behavior Genetics* **33**:591–606.

van den Oord, E. J. C. G., Verhulst, F. C., and Boomsma, D. I. (1996). A genetic study of maternal and paternal ratings of problem behaviors in three-year-old twins. *J. Abnorm. Psychol.* **105**:349–356.

van der Valk, J. C., Verhulst, F. C., Stroet, T. M., and Boomsma, D. I. (1998). Quantitative genetic analysis of internalizing and externalizing problems in a large sample of 3-year-old twins. *Twin Res.* **1**:25–33.

van der Valk, J. C., van den Oord, E. J., Verhulst, F. C., and Boomsma, D. (2001). Using parental ratings to study the etiology of 3-year-old twins' problem behaviors: Different views or rater bias? *J. Child Psychol. Psychol.* **42**:921–931.

van der Valk, J. C., van den Oord, E. J. C. G., Verhulst, F. C., and Boomsma, D. I. (2003). Using common and unique parental views to study the etiology of 7-year-old twins' internalizing and externalizing problems. *Behav. Genet.* (in press).

Verhulst, F. C., and Koot, H. M. (1991). Longitudinal research in child and adolescent psychopathology. *J. Am. Acad. Child Adolesc. Psychiatry* **30**:361–368.

Vierikko, E., Pulkkinen, L., Kaprio, J., Viken, R., and Rose, R. J. (2003). Sex differences in genetic and environmental effects on aggression. *Aggress. Behav.* **29**:55–68.